Uncertainties in LCA (Subject Editor: Andreas Ciroth)

How to Obtain a Precise and Representative Estimate for Parameters in LCA*

A case study for the functional unit

Andreas Ciroth and Michael Srocka**

GreenDeltaTC GmbH, Raumerstrasse 7, 10437 Berlin, Germany

** **Corresponding author** (ciroth@greendeltatc.com)

DOI: http://dx.doi.org/10.1065/lca2007.06.345

Please cite this paper as: Ciroth A, Srocka M (2008): How to Obtain a Precise and Representative Estimate for Parameters in LCA. A case study for the functional unit. Int J LCA 13 (3) 265–277

Abstract

Background, Aim and Scope. Quite often there is need for precise and representative parameters in LCA studies. Probably the most relevant have direct influence on the functional unit, whose definition is crucial in the conduct of any LCA. Changes in the functional unit show directly in LCI and LCIA results. In comparative assertions, a bias in the functional unit may lead to a bias in the overall conclusions. Since quantitative data for the functional unit, such as geometric dimensions and specific weight, often vary, the question arises how to determine the functional unit, especially if a comparative assertion shall be representative for a region or market. Aim and scope of the study is to develop and apply methods for obtaining precise and representative estimates for the functional unit as one important parameter in an LCA study.

Materials and Methods. Statistical sampling is applied in order to get empirical estimates for the weight of yoghurt cups, as a typical parameter for the functional unit. We used a two-stage sampling design, with stratified sampling in the first stage and three different sampling designs in the second stage, namely stratified, clustered, and *a posteriori* sampling. Sampling designs are motivated and described. In a case study, they are each used to determine a representative weight for 150 g yoghurt cups in Berlin, at the point of sale and within a specific time. In the first sampling stage, food markets are randomly selected, while in the second stage, yoghurt cups in these food markets are sampled. The sampling methods are applicable due to newly available internet data. These data sources and their shortcomings are described.

Results. The random sampling procedure yields representative estimates, which are compared to figures for market leaders, i.e. yoghurt cups with very high occurrence in the supermarkets. While single types of yoghurt cups showed moderate uncertainty, representative estimates were highly precise.

Discussion results show, for one, the performance of the applied statistical estimation procedures, and they show further that adding more information in the estimation procedure (on the shape of the cup, on the type of plastic, on the specific brand) helps reducing uncertainty.

Conclusions. As conclusions, estimates and their uncertainty depend on the measurement procedure in a sensitive manner; any uncertainty information should be coupled with information on the measurement procedure, and it is recommended to use statistical sampling in order to reduce uncertainty for important parameters of an LCA study.

Recommendations and Perspectives. Results for market leaders differed considerably from representative estimates. This implies to not use market leader data, or data with a high market share, as substitute for representative data in LCA studies. Statistical sampling has been barely used for Life Cycle Assessment. It turned out to be a feasible means for obtaining highly precise and representative estimates for the weight of yoghurt cups in the case study, based on empirical analysis. Further research is recommended in order to detect which parameters should best be investigated in LCA case studies; which data sources are available and recommended, and which sampling designs are appropriate for different application cases.

Keywords: Berlin; empirical data sampling; functional unit; representativeness; sampling design; statistical sampling; stratified sampling; uncertainty; yoghurt cups

1 Representative Data

Life Cycle Assessments need representative data for drawing well-founded conclusions about their object of study. For inventory datasets, several industrial branches in the EU have undertaken the task to provide data sets for 'their' processes that are representative for the industrial branch. Some LCA studies try to get representative datasets by considering large samples, i.e. high market shares; sometimes, high market share is quoted as indicator for good representativeness[1].

In order to obtain representative inventory data and conclusions, surveys [IAI 2003, FEFCO 2003, Boustead 2003] and expert judgement are often applied; the use of statistical sampling is not yet reported.

Statistical sampling and measurement, however, is a means to obtain truly representative data. The effort is often comparable to other approaches, or even lower. A classical example for the superiority of statistical sampling dates from the US presidential elections in 1936. In order to estimate who would win, two 'sampling methods' were performed, independently. Dr. Gallup conducted 3,000 interviews, with interviewees selected via (random) quota sampling of eligible voters, and predicted from these a victory of President Roosevelt over Landon, a republican, with 54% vs. 46%.

* **ESS-Submission Editor:** Seungdo Kim, PhD (kimseun@msu.edu)

[1] "Data for the Life Cycle Survey were obtained from: 82 world-wide aluminium electrolysis plants producing 14.7 million metric tons of primary aluminium, representing about 60% of world-wide aluminium smelting operations (base: primary aluminium from WBMS 24,464,400 t)." [IAI 2003, p. 4]

© Springer-Verlag 2008

The 'Literary Digest', a popular magazine at the time, polled 10 million people selected from automobile registration lists and telephone directories, evaluated 2.4 million surveys, and predicted a victory for Landon (57%). Actually, Roosevelt won the election by 61%! 3,000 interviews yielded a better result than 2.4 million surveys; seemingly the smaller sample reflected the eligible voters to a higher extent (consider that in 1936, telephones and cars were still predominantly owned by the middle and upper classes).

In Life Cycle Assessments, statistical sampling is barely applied[2]; it appears often impractical. One cannot force, for example, a company to provide process data merely because this company was selected in the sampling procedure. However, not all aspects of an inventory are critical for its representativeness. There are several options in sampling design which allow tailoring the design to meet specific challenges regarding data availability. And not least, availability of data has increased dramatically in the past months, due to legislation and business activities (e.g. [1–2,4]).

Among all the different parameters of an inventory system, the functional unit (f.u.) is often of prime importance [Kim Dale 2006, Hischier Reichart 2003, Cooper 2003]. It is like a pivot for the whole LCA model. Changes in the f.u. are directly reflected in LCA results, so, e.g., in a linear LCA model a 10% increase in f.u. means a 10% increase in calculated environmental impact. In comparative assertions, even the ranking between alternatives might change if f.u.'s of the compared systems do not change in a corresponding manner.

Any LCA that strives for representativeness thus should strive for representative data for the functional unit. The remainder of the article describes, for yoghurt cups, a procedure to obtain truly representative data by statistical sampling.

2 Background: Statistical sampling theory in brief

2.1 The idea of statistical sampling

Statistical sampling has developed into a credited field of science with broad practical experiences. Today it is applied in product testing, in political surveys, as well as in health science. This paper has neither room for a full introduction into sampling theory nor do the authors feel this would be their prime task. Instead, we provide some basics as necessary for the understanding of terminology and concept of the article. For interested readers, Cochran [Cochran 1977], and also [Sudman 1976], [Thompson 2002], and [Schwarz 1975] are good introductions.

Statistical sampling is a technique to collect data when little resources are available or when the overall population[3] is not completely accessible in an efficient and reliable manner [Cochran 1977, p. 3]). Sampling means, basically, 'drawing' a number of n elements from a population of amount N, n<N. Depending on the composition of the population, different approaches are used to obtain representative samples.

A sample is considered representative if and insofar its composition 'represents', or reflects, that of the population. If the sampling procedure is designed in a satisfying manner, results can be as good as if the whole population was analysed. In some cases, results are even better as the diligence for a single item can be enhanced. Any figures of interest are calculated from the sample and generalised in order to obtain figures for the target population by estimating them from the sample. Typically, these figures are 'average' and 'spread' of numerical properties of an object of study.

As an example, the object of study could be yoghurt cups of 150 g yoghurt, and the calculated parameters the average weight and the standard deviation (or variance) of the weight of the cups, for the targeted population.

A sound design of the sampling procedure and an adequate estimation function is essential for the quality of the estimate. A first step is to clearly define the population, which should, in turn, be determined by goal and scope of the study.

2.2 Sampling designs

There are several choices in sampling design; properties of the population, the characteristics to be sampled, and, not least, the desired sampling precision and affordable costs and effort, will determine the sampling design that fits best for a study.

Simple random sampling is the easiest sampling design. From a population with N elements n elements are selected randomly without replacement so that each element in the population has the same chance of being drawn. In many cases, a further division of the population reduces the variance of the estimate and will thus improve its quality. If the population is rather heterogeneous and may be split into subsets which are more homogenous, each, the variance from a simple random sample of the whole population will be higher than necessary. More homogenous subsets will have a lower variance, respectively. Now, if the subsets can be assumed to be independent, which is often the case, then the overall variance is simply the sum of the variances of the subsets. In an extreme case, with completely homogenous subsets, their variance is zero, and the resulting overall variance is calculated to zero as well. The subsets are called strata, one subset a stratum. Overall the 'stratification' has a positive effect if the calculated variance is much lower than if calculated directly, without stratification. In **stratified random sampling**, one will take a simple random sample from each stratum [Cochran 1977, pp. 89].

Imagine yoghurt cups of different shape, broad and narrow ones. The shape will influence the weight of the cup; it can be observed easily and thus suits as an indicator for the definition of homogenous and disjoint (independent) strata. Calculating the mean of the weight within these strata of cups of the same shape will produce a precise estimate, for each shape, with only a few measurements in each stratum. If the shares

[2] An exception is the LIME project: [Itsubo Inaba 2003] applied sampling in a survey for the representative Japanese population for monetarising environmental endpoints.

[3] 'population' is the term used in statistics to describe the set where the sample is taken from [Sudman1976, p. 11]; in a narrow sense, this population is called the sampled population. The sampled population should match the population about which information is wanted, the population that is targeted, the target population. [Cochran 1977, p. 5]. Since a sampling study attempts to generalise results from the sample to the target population, the latter is also called 'the universe' [Sudman 1976, p. 12].

of cups of different shapes in the overall population are known, then a more precise mean for the population can be calculated by combining the estimated means from the strata.

A drawback is that in order to know the shares (of each type of cup, e.g.) and to be able to draw from each stratum, a reliable and complete list of the elements and their attributes in the population of interest is needed, which is often not available, or at least time and cost consuming to obtain. For example, one would need a complete list of yoghurt cups and their shapes in all supermarkets in Berlin.

In contrast, a list of 'higher-level entities' (a list of supermarkets, for this case) is often available. This calls for **cluster sampling**, a different sampling design. The idea is to group the elements in the population in classes, which are called clusters, and then perform a random sampling on these clusters and determine all elements in the selected clusters in a second step. In the example above, one would first perform random sampling on the 'supermarket list' and afterwards determine all the yoghurt cups for the selected supermarkets. Quite contrary to stratified sampling, where the strata have to be as different as possible and the elements within the strata as similar as possible, the clusters in cluster sampling have to be as similar as possible and the elements within the clusters as different as possible.

And if the property of interest (the weight of a yoghurt cup) is obtained only with some effort, sampling on large populations might become tedious. In this case, a second sampling step (that 'samples from the sample') is often convenient. This type of design is called a **two-stage sampling**

The sampling techniques described here represent only a small extract of possible sampling designs. Depending on the task, other designs and combinations might be adequate. For a description of further designs, with emphasis on variance estimation, see e.g. [Wolter 1985].

3 A Case Study: Uncertainty in the weight of yoghurt cups in Berlin

3.1 Aim of the exercise

Aim of the case study is to provide a precise and representative estimate for the weight of plastic yoghurt cups for 150 g yoghurt at point of sale in Berlin, in food markets, available for consumers on one specific day. The precision shall be expressed as standard deviation. In an exploratory data analysis step, uncertainty for different types of yoghurt cups shall be determined. The sampling design needs to take into account how yoghurt cups are presented to customers in Berlin. Different designs shall be evaluated.

3.2 Inclusion of the marginal consumption in a comparative LCA

The population are all yoghurt cups that contain 150 g yoghurt which are presented to consumers in supermarkets and other food markets in Berlin, on a specific day. The type of yoghurt (natural yoghurt, with fruit) may vary. The cups must be made from plastics; compound material (carton, paper, and plastic) is excluded from the analysis.

One type of yoghurt in one market is considered as one element in the population, several cups of the same type in one market are not distinguished, while the same type in another market is another element in the population.

The number of the different types of yoghurt cups in different markets, in other words the number of elements N in the population, is not known beforehand.

Thus applying a two-stage sampling design is reasonable. In the first stage, the population is the number of supermarkets and food markets in Berlin. From this population, which will be called 'market population' in the following, representative markets are selected by stratified random sampling, and the number of different types of yoghurt cups is determined within this sample, the 'market sample'. This is the first sampling step.

From the market sample, an estimator is available for the number of different yoghurt cup types in Berlin. This is another population, which will be called 'cup population'. Its elements are the different types of yoghurt cups available in Berlin at point of sale, and at the time of sampling. For the selected markets, both cluster sampling and stratified sampling was applied. An a posteriori stratification was applied to the cluster sampling in order to investigate further possible improvements of the estimator (**Fig. 1**).

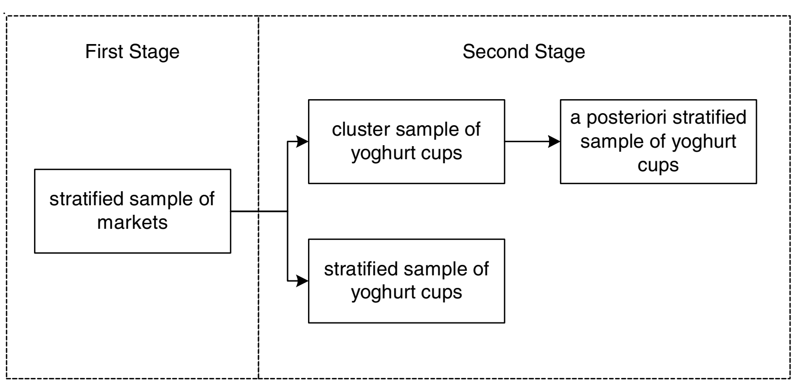

Fig. 1: Sampling designs applied in the case study

3.3 Sampling procedure

In the first stage of sampling, for the market population, it is assumed that the type of yoghurt cups is similar for supermarkets of the same brand. Accordingly, the brand of the supermarket is used as criterion for defining the market strata. For larger brands, each brand is one stratum. Smaller brands with less than 40 markets were grouped into one stratum.

The number of different markets in Berlin is easily available via internet research. Yellow pages [1] and Google Maps [2] provided, in combination, an overview of markets and their brands and location. Overall 1215 markets were found and compiled into a list. The strata and the aggregated number of markets are given in **Table 1**. Related brands that belong to a single company but are intended for different types of customers (e.g. budget-priced vs. regular range of goods) were treated as different brands.

Based on this list, stratified random sampling was performed with units proportional to size; the probability of a market to be drawn being proportional to the number of markets of this brand. The sampling ratio was set to 3% (3% of each stratum was drawn), hence elements from larger strata are more frequent in the sample. Overall, 35 markets were randomly selected.

Table 1: Number of food markets in Berlin, per market brand

Market stratum	Market	Number of markets in Berlin city
1	Aldi	183
2	Kaiser's	173
3	Plus	162
4	LIDL	93
5	SPAR	86
6	Penny	72
7	EDEKA	71
8	Rewe	63
8	miniMal	31
9	Reichelt	54
10	Netto	43
11	Meyer&Beck	41
12	extra	35
12	Butter-Lindner	31
12	Kaufland	17
12	Norma	17
12	Karstadt	13
12	Real	10
12	Bio Company	9
12	Kaufhof	5
12	Ullrich	3
12	Birlik	1
12	K-Markt	1
12	LPG BIO Markt	1
Σ		**1215**

The selected markets were visited and the yoghurt cups in these markets inspected, noting properties of all yoghurt cups that belonged to the population (excluding e.g. cups made from cardboard). Four of the drawn markets were found to be out of business. Three markets were added as replacements.

For the cluster sample sixteen markets of the first stage were randomly selected and all of the respective yoghurt cups in these markets were analysed. Markets in the cluster sample were located off-centre, making them more difficult to visit. If the location of a market had any influence on the weight of the yoghurt cups, a bias in the cluster sample would be introduced. This was assumed to be not the case, which was tested later on. The following attributes were thought to influence the yoghurt cup weight in the population: (i) type of plastic (PP, PS); (ii) colour and opacity (clear, white); (iii) width (small, wide), and (iv) overall shape (cups with foot and without foot).

These attributes were used for defining the strata which resulted in 16 different ones for the stratified yoghurt cup sample. **Table 2** shows the strata and the 'assigned' yoghurt cups together with the characteristics of each stratum.

In a next step, random sampling was performed, per stratum, with a sampling ratio of 20%. A yoghurt cup in one stratum thus has a chance of 20% to be drawn.

Fig. 2 shows the markets in the sample, the number of different types of 150g yoghurt cups offered to customers in each market, and the cups drawn for the stratified sample and the cluster sample.

All of the respective cups in cluster and stratified sample were bought, emptied, any labels were carefully removed, and the empty cups were weighed on a high precision, laboratory balance[4]. Data was entered into an MS ACCESS database, checked for errors, and further analyses were performed by using the statistical open source software package R (Version 2.3.1) [3] and spreadsheet software.

3.4 Uncertainty in sampled data

The following figures show the uncertainty for all yoghurt cups, for two selected types of yoghurt cups, for all sampled clusters, for all sampled strata, and, finally, a comparison of the uncertainty for the overall sample, for one cluster, for one stratum, and for one type of yoghurt cup.

For all types of yoghurt cups, the weight varies between less than four and more than eight grams. The shape of the histogram resembles, slightly, a log normal probability distribution, with two peaks (maximum frequency) at around 5 and 6 g which might indicate an overlap of two different distributions[5]. The overall sample is a blend of the weights of many

[4] Weighing machine: OMNILAB OL 210-A, max 220 g ± 0.0001 g, from NovaBiotec laboratory, Berlin.

[5] A histogram is a graphical representation of tabulated frequencies; it is often used to explore the shape of data distributions.

Table 2: Yoghurt cup strata in the sample, their characteristics and respective yoghurt brand names (note that the stratum nr. is different from the one shown in table 1; PS: Polystyrene)

Stratum no.	PS?	Clear Plastic?	Narrow?	Foot?	Yoghurt brand names (in some cases description of shape is added)
6	Yes	No	Yes	No	Alpa Frucht Yoghurt; Alpa Joghurt pur 4*150; Berchtesgadener Land Bioghurt; Biac probiotischer Joghurt 4*150; Bio Wertkost Bio Joghurt; BioBio Fruchtjoghurt; BioBio Naturjoghurt; Bissou 4*150; demeter Biogurth; Ehrmann Almighurt; Ehrmann Genuss Diät; Fruchtjoghurt mild 4*150; Fruchtjoghurt mild 8*150; Grazil Fruchtjoghurt 4*150; Herzgut Fruchtjoghurt; Martinshof Ziegenjoghurt; Mibell Joghurt PS; Milram Magermilch Joghurt; Mondelice Fruit Split 4*150; Müller Joghurtschnee; Naturell 3.5% Fett; nöm l.free; Rogge Bio Joghurt; Sachsenmilch Früchtchen; Söbbeke; Yogosan Fruchtjoghurt 4*150; Bissou Joghurt auf Frucht
8	Yes	No	No	No	Grazil Joghurt; Sachsenmilch Vanillezauber; Yoganic 0; yes wide
9	No	Yes	Yes	Yes	Elite Joghurt mild (Fuß)
10	No	Yes	Yes	No	Campina Fruchtstrudel; Dr. Oetker Jobst; Elite Vanilla auf Frucht (clear); Gut & fein Joghurt auf Frucht; ja! Joghurt mild (clear); Müller Froop
13	No	No	Yes	Yes	Elite Sahnejoghurt mild (foot); Gut & Günstig Sahnejoghurt; Mertinger Sahnejoghurt; Sahnejoghurt; Tipp Sahne Joghurt mild; Zott Sahnejoghurt
14	No	No	Yes	No	Bauer Joghurt mild; Bioness Bio Fruchtjoghurt; Campina Milchreiter; Campina Optiwell; Elite Joghurt mild; Elite Karamell/Vanilla; Elite Vit Balance; Erlenhof Joghut mild; Goldblume Joghurt mild; Gropper probiotic Jogurtcreme; Gut & fein Joghurt mild; Gut & fein Joghurt natur 4*150 ; ja! Joghurt mild; ja! Magermilchjoghurt; Mark Brandenburg Joghurt mild; Mibell Joghurt; Milbona Sahnejoghurt; Pro Jogo 4*150; proactiv diät; Tipp Joghurt Creme; Tipp Joghurt mild; Yoganic 0.1% (narrow); Zott Jogole
16	No	No	No	No	Bauer Yogorande; Choco Picnic Joghurt + Schokoraspel; Elinas Joghurt nach griechischer Art 4*150; Hofmaier Cream-Jogh.; Hofmaier Joghurt & Chocosplits; Landliebe Joghurt; Minus L Joghurt; Mövenpick Feinjoghurt; Yocous; Yogosan Edelrahmjoghurt

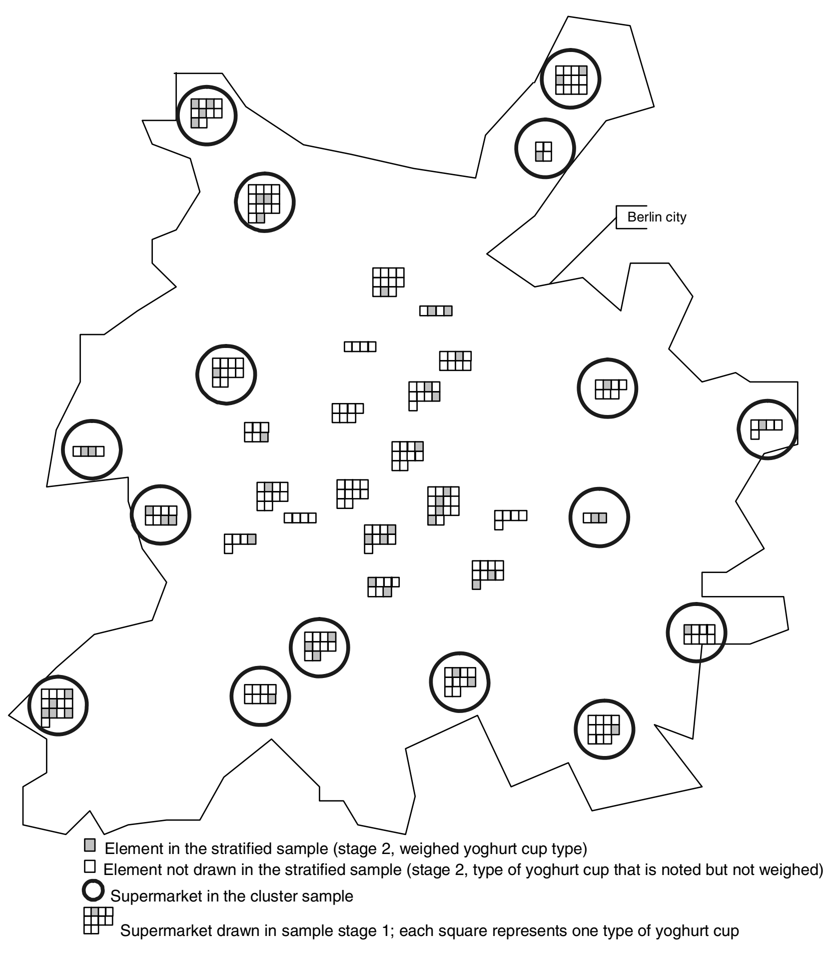

Fig. 2: Location of the supermarkets in the sample, the number of yoghurt cups in each market (squares), and the yoghurt cups in the stratified sample (greyed squares) and the cluster sample (circles)

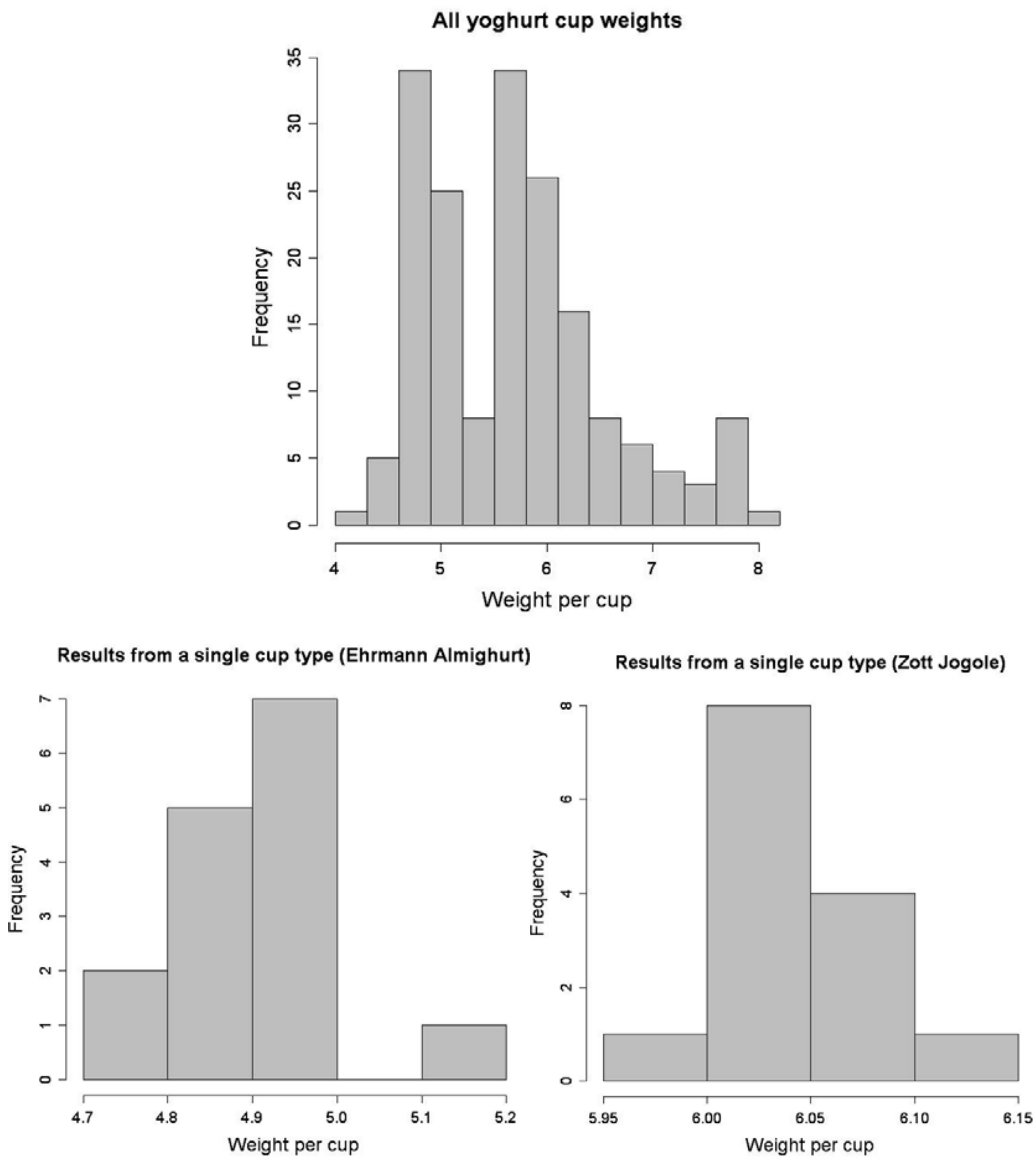

Fig. 3: Histograms for the overall sample and for two different yoghurt cups, provided as examples

different types of yoghurt cups. For each type, the variation in weight is lower than for the overall sample and the histograms often resemble normal or lognormal probability distributions. In **Fig. 3**, this is shown for two specific cup types in comparison to the histogram of the overall sample.

The coefficient of variation, calculated as the ratio of empirical standard deviation to mean, per type of yoghurt cup, lies in the range from almost zero to eight percent, with the majority between one and three percent (**Fig. 4**, see Annex, p. 277). This can be interpreted as the relative uncertainty in the sampled data.

The cluster sample is shown in box plots in **Fig. 5**, one cluster being one supermarket. Two clusters (the two on the left side of the figure) have quite narrow samples; these are two discount supermarkets which offer only a few different types of yoghurts. The other clusters have a high variation per cluster, but seem rather similar regarding quartile range, and to some extent also regarding to the mean (the thick black line in each box). The strata, on the other side, clearly differ from one to another, and have a much lower variation per stratum (see Fig. 5). Uncertainty in strata and clusters thus reflects the aim of the sampling (homogenous strata, cluster according to the population, see section 2).

3.5 Calculation procedure for a representative estimate

Formulas for the calculation of estimates in stage 1 and stage 2, for stratified, cluster and a posteriori stratification as well as for calculating uncertainty are given in the Annex, p. 276.

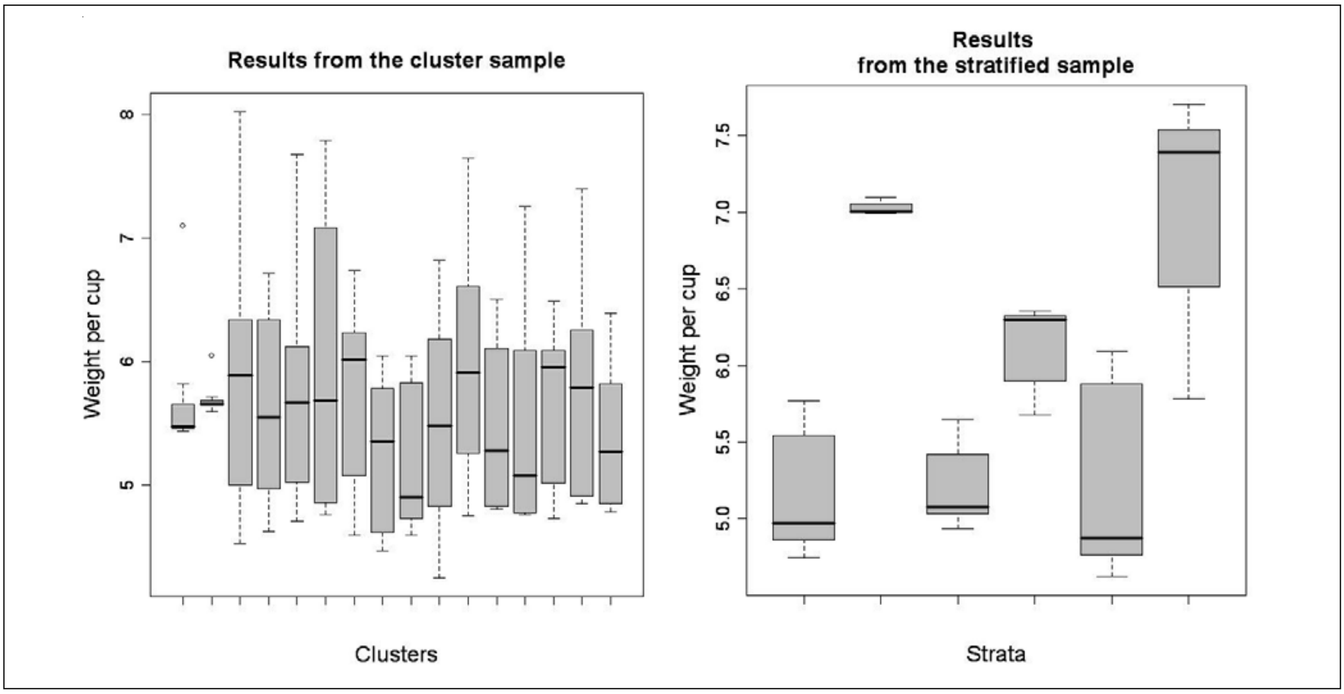

Fig. 5: Boxplots for the cluster sample, per cluster, and for the stratified sample, per stratum

3.6 Results

The following results are achieved with calculation formulas as described in section 3.5. They refer to the estimated mean and variance of the weight of one 150 g yoghurt cup in Berlin, representative for the day of sampling:

Stratified sampling:

$\hat{\hat{Y}} = \sum_{h=1}^{M} \frac{N_h}{N} \bar{y}_h = 5.6474\,\text{g}$, with a variance of

$$\hat{Var}(\hat{\hat{Y}}) = \sum_{h=1}^{M} \left(\frac{N_h}{N}\right)^2 \frac{N_h - n_h}{N_h} \frac{s^2{}_h}{n_h} = 0.0034$$

With:

$\hat{\hat{Y}}$ = Estimate for the mean
h = Index number of stratum
N = Number of food markets in Berlin (population 1)
N_h = Number of food markets of stratum h in Berlin
M = Number of strata
n_h = Number of food markets of stratum h in the sample
$\bar{y}_h$ = Average number of types of yoghurt cups in a food market in stratum h
s_h^2 = Variance of the number of types of yoghurt cups in stratum h

The uncertainty in the population is calculated to $s^2 = 0.72$, the variation coefficient to 15%.
Clustered sampling:

$\hat{\hat{Y}}_{CL,QS} = 5.6589\,\text{g}$, with a variance of $\hat{Var}(\hat{\hat{Y}}_{CL,QS}) = 0.0072$.

With: $\hat{\hat{Y}}_{CL,QS}$ Estimate for the mean, clustered sampling

The variance is equivalent to a standard deviation of 0.085, and to a coefficient of variation (relative uncertainty, standard deviation divided by the mean) of 0.015 or 1.5%.

A posteriori stratification of the cluster sample:

$\hat{\hat{Y}} = 5.7245g$ and $\hat{Var}(\hat{\hat{Y}}) = 0.0014$.

With: $\hat{\hat{Y}}$ Estimate for the mean

The estimated variance is even lower than in the cluster sampling; it corresponds to a standard deviation of 0.037, and to a variation coefficient of 0.0065 or 0.65%.

3.7 Discussion

3.7.1 Weaknesses of the study

Several weaknesses of the study are worth being mentioned.

The weight of the cups could be influenced by remaining yoghurt in the cup. It could also be influenced by water adherent either to the walls of the cup or to paper etiquettes. We aimed to minimise these effects by thoroughly cleaning the cups, by drying them at room temperature, and by removing any paper etiquettes. Cups made from a combination of paper and plastic were excluded, see section 3.2.

Some shops from the original sample were closed. In the sampling procedure, they were replaced by other similar markets. Obviously non-existent markets cannot be discovered for the overall sample, i.e. for those markets that have not been visited. In the sample, the share of closed or replaced markets was about 10% (4 of 35). This figure may influence the result (the estimates) if 'new' markets had a

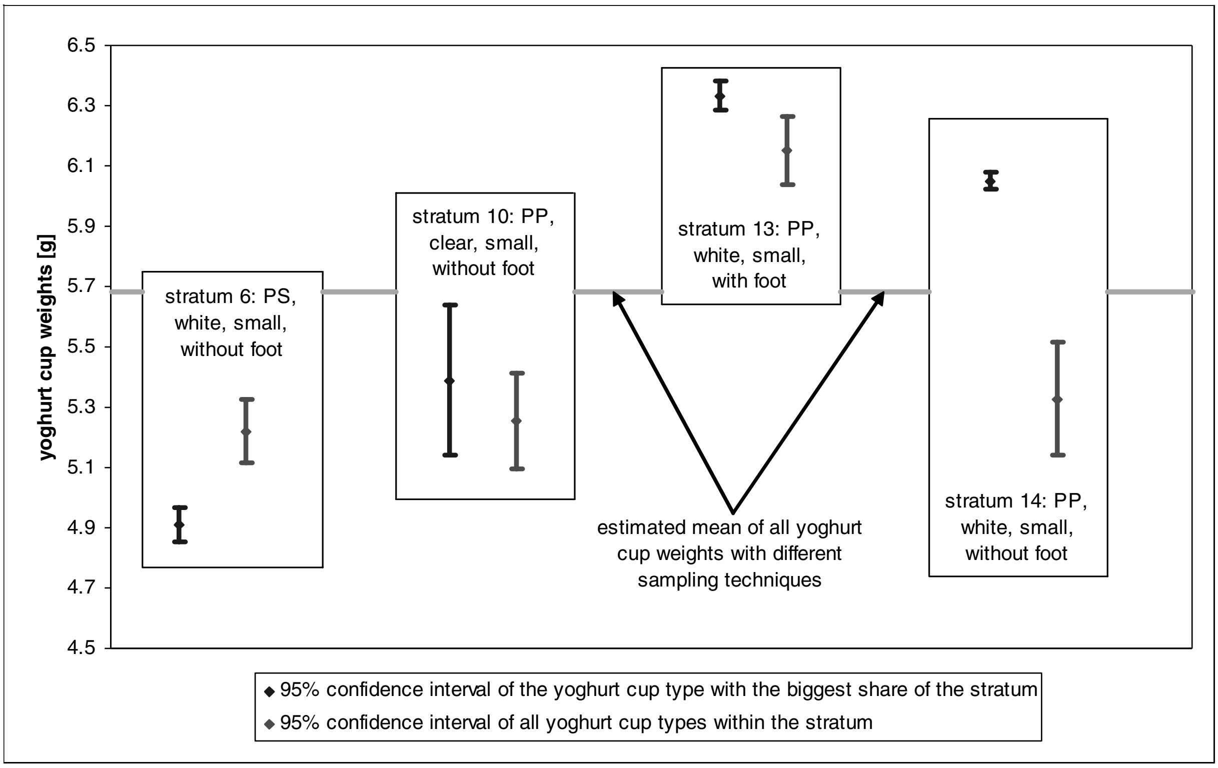

Fig. 7: Range of yoghurt cup weights in the sample, and estimated means of the three sampling techniques. Confidence intervals for the estimated mean are too small to be visible in the graphic

different yoghurt product portfolio from markets that left business. We have, at the moment, no such indication. In any case a 10% rate is a sign that the basic data does not match the speed of market transformation.

In the shops, the cups were selected 'one per type', without specific sampling pattern for each type. Individual sampling patterns might have existed and could lead to a biased selection of cups.

3.7.2 Sampling design

In order to analyse the effects of the sampling design, cluster sample and the stratified sample are tested in a Chi-Square test[6]. Null hypothesis for the test is: The weight of the yoghurt cups in the stratified sampling obeys the probability distribution of the weight of the yoghurt cups in the cluster sample.

As a result, with a significance level of 5%, and 5 classes in the test, the test statistic X^2 is calculated to $(X^2 = 1.3572) < (X^2\,(0.95,4) = 9.4877)$. If this inequality was not fulfilled, the null hypothesis would need to be rejected, which is clearly not the case here. Hence, it seems fair to assume that both stratified sampling and cluster sampling have identical probability distributions. For further illustration, **Fig. 6** (see Annex, p. 277) shows a histogram of cluster and stratified sampling. Both distributions have about the same shape which supports the test results. One can thus assume that cluster samples and population follow the same probability distribution, which is also a confirmation of the quality of the cluster samples.

On a more intuitive basis, **Fig. 7** shows the overall estimator for the mean, the confidence intervals within four different strata, and the confidence interval for the type of yoghurt cup with the highest share in each shown stratum. Obviously, results per stratum differ considerably, and often cups with the highest share differ considerably from the stratum where they belong to (stratum 6 and 14). In contrast to this, the estimated representative mean is for all sampling methods similar to an extent that its differences do not show in the figure, and the estimator is precise to an extent that the confidence interval does not show, either.

Fig. 7 shows also that all sampling techniques lead to very similar results. For each of the designs applied, the result is very precise, with a very small variance leading to very small confidence intervals.

3.7.3 Representativeness

A demand for representative results is often intended to be met by expert judgement alone, or by using parameter values that represent a high market share, instead of applying statistical sampling. While expert judgement seems rather a

[6] A Chi-Square (X^2) test is a standard test for comparing two different value series, see e.g. [Sachs 1992, pp. 420]; tests are usually conducted by stating a null hypothesis ('these two data samples have the same distribution'; 'all dogs are black bulldogs'. Then, a test criterion (or several criteria) are determined to test whether this hypothesis can be accepted or not, based on available (sample) data. For details please refer to standard statistic text books, e.g. [Tabachnik Fidell 2006].

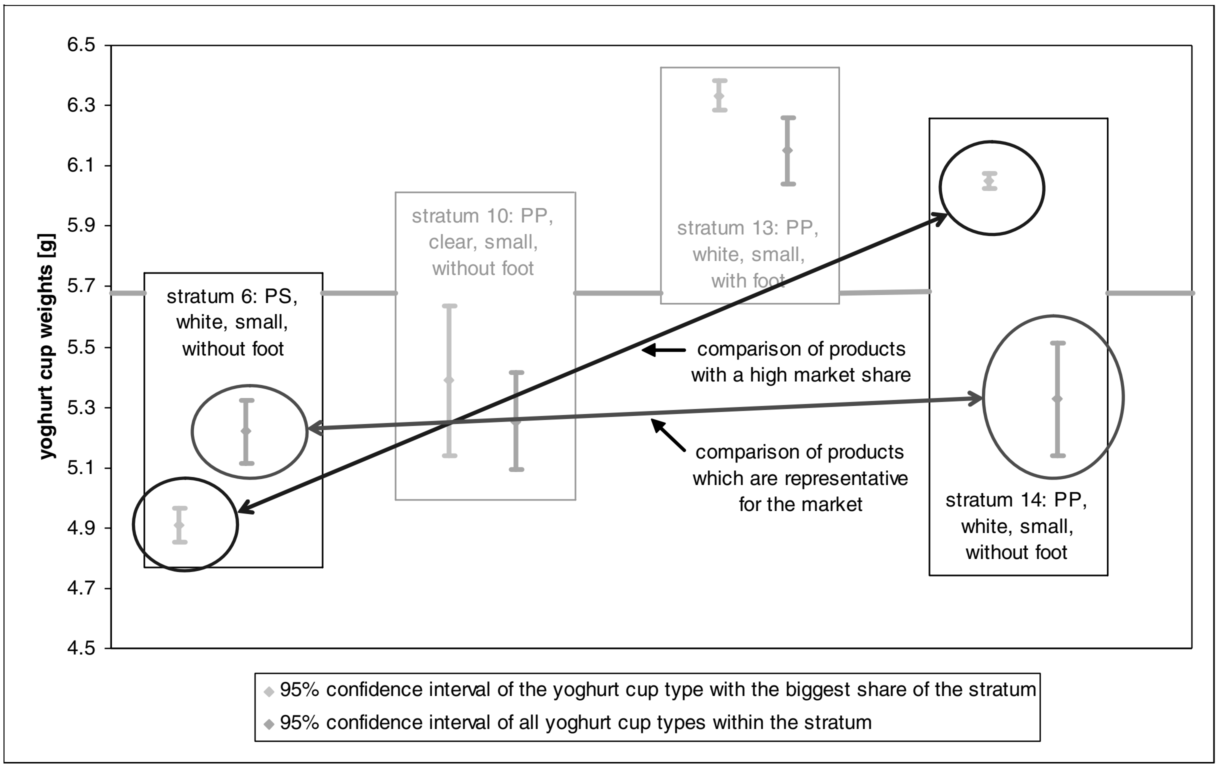

Fig. 8: Comparison of the market leaders and of representative products

fall back option (used if no further information is available), the reason of high market share deserves attention.

Fig. 7 shows how different the mean of the population (here: one stratum as a set of yoghurt cups with identical shape) may be from the mean of its most common elements.

This bias may have severe consequences in a comparative assertion, which is demonstrated in **Fig. 8**. Stratum 6 and stratum 14 (white plastic, narrow shape, without foot) differ only in the type of plastic (PP vs. PS). The most common yoghurt cups, the 'market leaders', have a share of 31 and 30% in the respective stratum. In the case of PS, the market leader is much lighter than the average, while for PP, it is, on the contrary, much heavier. The average weight, and even the confidence intervals, for both strata lie in the same range. Thus, a comparison based on market leaders, or a high market share, will produce, in this case, a completely different picture than a comparison based on representative samples. The difference is considerably high, 4.9 to 6.1 g for the market leaders, or about 25%. Note that these are differences in the weight of the functional unit which will, in a linear, common LCA models, directly show in the result. Considering only the products with the highest market share would thus lead to a fatally wrong conclusion[7].

[7] Case study results in this paper reflect the market share per different types of cups offered to customers, and not per unit sold. Since some cups are offered in packed units of four or eight, both will be different, and thus should be discerned. However, it seems likely that the principle discovered on the basis of the empirical data in this paper, namely that market leaders might differ from the rest of the market in study-relevant aspects, will hold also for a market share that is based on the number of units sold.

Statistical sampling approaches offer a possibility to come to representative data; they, on the other hand, demand a design tailored to meet data availability challenges in order to be applicable.

3.7.4 Uncertainty

The case study provides empirically based uncertainty information for the functional unit as one prime parameter in an LCA study.

When expressed as coefficient of variation, the uncertainty is 15% for the overall population; for single types of yoghurt cups it is in the area of one to three percent, with two outliers of about eight percent. For the estimated mean, this uncertainty can be reduced dramatically, to 0.07% for the stratified sampling. This is an effect of the sampling procedure, and of additional information taken into account by the sampling procedure (shape of the yoghurt cup, material, and so on).

For quantitative figures, the results indicate that an LCA on yoghurt packaging which considers the brand name will face a relative uncertainty (coefficient of variance) between one and three percent from the functional unit alone. This is, of course, not much. Is it then relevant? There is no easy answer to this question.

The functional unit is the first quantitative datum in an LCA calculation. Many other will follow. These 'following data' are of course also more or less uncertain. In the most basic case, an LCA product system is a linear chain of processes. A relative uncertainty of 2 percent in each product flow

yields, after 7 processes, an overall uncertainty (propagated uncertainty plus uncertainty in each flow) of about 15%[8].

In any case study, the impact of uncertainty[9] will depend on the structure of the product system; in comparative assertions, it will also depend on amount and location of uncertainties in the compared product system, and on its structure. Quite often, similar processes in both compared systems are omitted; this might either reduce sources of uncertainties but might also lead to a higher relative uncertainty in the result. These topics are not completely solved but have been discussed in earlier literature, e.g. [Ciroth 2001], [Ciroth et al. 2004], [Huijbregts 2001].

3.7.5 Similar approaches in marketing

In marketing, one often encounters the need to control the effect of a marketing campaign, or, more basically, to learn more about consumer behaviour, or about competitors. Quite often, yellow pages, web directories, and similar information is consulted in these cases ('the Yellow Pages as a market research tool', [Gross et al. 1993, p. 147]; 'Marketers should regularly and systematically utilize numerous sources for competitive evaluation', states Weinrauch, [Weinrauch 1987, pp. 18], listing about 25 different sources. While of course goal and scope of marketing is not fully in line with those of a Life Cycle Assessment, the need to analyse, carefully, the circumstances in which the product is used, and how the product exactly looks like at the point of sale, is very comparable between both disciplines. This is obviously only an analogy at first glance which needs to be explored more in detail, regarding the applied tools (for marketing, e.g. [Fitzroy 1976]) and their performance.

4 Conclusions and Perspectives

Yoghurt cups and other plastic food packaging have often been the subject of Life Cycle Assessments or similar analyses, e.g. [IFEU 2003], [Petcore 2004], [IFEU 2006], [Keoleian et al. 2001]. The functional unit, in each case, will often depend on the weight of the cup in a linear manner. In consequence, differences in the weight per packed volume in, e.g., five percent between two alternatives will lead, ceteris paribus, to a difference in impact assessment results of five percent as well (assuming again a linear LCA model). Thus there is a clear need for reliable and precise information for the parameter 'weight', especially in comparative assertions.

The sampling of yoghurt cups, in the case study, provided precise and representative estimates, based on empirical investigation. The effort was manageable. Three different, multistage sampling designs were tested; all of them provided precise estimates of the mean weight of a yoghurt cup. The cluster sampling had practical advantages in the case study.

The results clearly demonstrate that high market share does not at all ensure representativeness. Instead, products (yoghurt cups) with high market shares often were very different from a representative average. The results indicate, in addition, low uncertainties (standard deviations) for most cups, and moderate uncertainties for single types of yoghurts. Using more, available, information in the sampling procedure and in the calculation of estimates for the mean reduced the uncertainty in the estimates considerably, yielding highly precise estimates.

This demonstrates that more information reduces uncertainty, or, in other words, observed uncertainty is the result of changes in data that cannot be explained or 'understood' otherwise. A change in the shape of a yoghurt cup often influences its weight; if the shape is known, the uncertainty in the estimated weight will decrease. If the brand name is known in addition, this will, with a specific sampling plan, help to further reduce the uncertainty[10].

Further, this indicates that 'uncertainty' is not a specific, fixed characteristic for a yoghurt cup, as would be the type of plastic; rather, uncertainty of the weight depends on two main parameters:

First: What is the precise object of study, what is the 'population' in terms of statistical sampling? '150 g yoghurt cups' would be too unspecific for the case. The specific type of yoghurt cup, time, and geographical scope are relevant additional aspects;

Second: How does the measurement procedure look like? A smart sampling procedure can reduce uncertainty.

It seems fair to assume that these two parameters hold for all quantitative data in LCA studies. Somewhat surprisingly, both are fully covered by the definition of a functional unit according to ISO 14040[11]. However, the measurement procedure is rarely addressed in LCA literature (examples are the Cascade project, [Cascade 2003]). The measurement procedure for uncertainty is addressed even less often. One of the few examples is [Sugiyama et al. 2005] who discuss how to obtain probability distributions from survey data.

We conclude that the measurement procedure for uncertainty shall be mentioned if uncertainty information is provided with quantitative data in LCA. Further, we conclude that

[8] $(100\% + 2\ \%)^7 \cong 115\%$; for the simplest case, assuming that uncertainties are independent.

[9] 'Impact of uncertainty' meaning here: how will the quantitative results, and ranking, and conclusions drawn from the study be affected from uncertainty.

[10] This interpretation is in line with classical information theory by Shannon: Shannon defines with H the information entropy, which is "a measure for […] how uncertain we are at the outcome [of selecting several events]" [Shannon 1948, p. 10]. For two events x, y holds: The information entropy of x and y is equal or lower than the entropy for y alone. Shannon: "The uncertainty of y is never increased by knowledge of x. It will be decreased unless x and y are independent events, in which case it is not changed". If we define measuring weight and other properties of yoghurt cups as events, then, according to Shannon, the uncertainty of the weight will decrease by knowledge of another property of the yoghurt cup unless the property does not influence the weight.

[11] While this statement is evident for the specification of the object of study, it is also true for the specification of the measurement procedure. ISO 14040, section 5.2.2, Function, functional unit and reference flows: "The functional unit defines the quantification of the identified functions […] of the product". Quantification implies the measurement procedure and the quantified measurement value.

analysing market leaders alone is not justified when looking for representative data.

Finally, it seems fair to state that more research is needed on appropriate sampling procedures, data sources, and parameters of LCA product systems which should undergo statistical sampling. Emerging new data sources, with acceptable accuracy, will foster the application of statistical sampling methods which are, when applicable, clearly superior to expert judgement and to analysing market leaders alone.

Acknowledgement: This study was supported by NovaBiotec GmbH, Berlin. We thank Jutta Hildenbrand of the University of Wuppertal for comments on an earlier version.

References

[1] http://www.meinestadt.de/berlin/branchenbuch, website August 2006

[2] http://maps.google.de/ , website August 2006

[3] The R Project for Statistical Computing, http://www.r-project.org/

[4] http://portalu.de/ website August 2006

[Bahr 2004]: von Bahr B, Steen B: Reducing epistemological uncertainty in life cycle inventory, Journal of Cleaner Production, Volume 12, Number 4, May 2004, pp 369–388(20)

[Boustead 2003]: Boustead, I: Eco-Profiles of the European plastics industry, Methodology, A report for APME, Brussels, July 2003

[Cascade 2003]: Weidema B, Masoni M, Cappellaro F, Carlson R, Notten P, Pålsson A, Patyk A, Regalini E, Sacchetto F, Scalbi S: Procedural guideline for collection, treatment, and quality documentation of LCA data; Task 2.3 of the Cascade project on Standards or Modelling LCA data, EU Contract No G7RT-CT-2001-05045

[Ciroth 2001]: Ciroth A: Fehlerrechnung in Ökobilanzen, Dissertation TU Berlin 2001

[Ciroth et al. 2004]: Ciroth A, Fleischer G, Steinbach J: Uncertainty Calculation in Life Cycle Assessments - A Combined Model of Simulation and Approximation, Int J LCA 9 (4) 216–226 (2004)

[Ciroth Srocka 2006]: Ciroth, A, Srocka, M: Uncertainty quantification for the functional unit, poster, Setac Annual Meeting, Den Haag 2006

[Cochran 1977]: Cochran WG: Sampling Techniques, 3rd edition, John Wiley, New York 1977

[Cooper 2003]: Cooper JS: Specifying Functional Units and Reference Flows for Comparable Alternatives, 8 LCA (6) 337–349 (2003)

[Edwards 1999]: Edwards PN: Global climate science, uncertainty and politics: Data-laden models, model-filtered data: Science as Culture, vol 8, no 4, pp 437–472, 1999

[FEFCO 2003]: FEFCO (Europ. Federation of Corrugated Board Manufacturers), Groupement Ondulé, European Containerboard Organisation (edts): European Database for Corrugated Board Life Cycle Studies, 2003

[Fitzroy 1976]: Fitzroy PT: Analytical methods for marketing management, McGraw-Hill, London 1976

[Gross et al. 1993]: Gross AC, Banting PM, Meredith LN, Ford ID: Business Marketing, Houghton Mifflin, Boston 1993

[Hischier Reichart 2003]: Hischier R, Reichart I: Multifunctional Electronic Media – Traditional Media. The Problem of an Adequate Functional Unit: A case study of a printed newspaper, an internet newspaper and a TV broadcast, 8 LCA (4) 201–208 (2003)

[Huijbregts 2001]: Huijbregts MAJ: Uncertainty and variability in environmental life-cycle assessment, academisch proefschrift, University of Amsterdam, 2001

[IAI 2003]: International Aluminium Institute (IAI): Life Cycle Assessment of Aluminium: Inventory Data for the Worldwide Primary Aluminium Industry, Mach 2003

[IFEU 2003]: IFEU GmbH: Ökobilanz Kunststoffbecher für Milchmischgetränke, Andreas Detzel, personal communication, 2003

[IFEU 2006]: IFEU GmbH: Life Cycle Assessment of Polylactide (PLA), Heidelberg, 2006

[Itsubo Inaba 2003]: Itsubo N, Inaba A: A New LCIA Method: LIME has been completed, Int J LCA 8 (5) 305 (2003)

[Kauermann 2006]: Kauermann G: Stichprobenverfahren; VL Skript, Universität Bielefeld, 05/02/2006

[Keoleian et al. 2001]: Keoleian GA, Brachfeld D, Dritz T, Kodama S, Phipps A, Steiner E: Life Cycle Assessment of the Stonyfield Farm Product Delivery System, The Center for Sustainable Systems, University of Michigan, Ann Arbor, Michigan, 2001

[Kim Dale 2006]: Kim S, Dale B: Ethanol Fuels: E10 or E85 – Life Cycle Perspectives, 11 LCA (2) 117–121 (2006)

[Petcore 2004]: Petcore, Brüssel: Ökobilanz für PET-Einwegsysteme unter Berücksichtigung der Sekundärprodukte, Endbericht, August 2004

[Sachs 1992]: Sachs, Angewandte Statistik, Springer Verlag Berlin Heidelberg 1992

[Schwarz 1975]: Schwarz H. Stichprobenverfahren, Oldenbourg, München 1975

[Shannon 1948]: Shannon CE: A Mathematical Theory of Communication, The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, July and October 1948

[Sudman1976]: Sudman S: Applied Sampling, Academic Press, New York 1976

[Sugiyama et al. 2005]: Sugiyama H, Fukushima Y, Hirao M, Hellweg St, Hungerbühler K: Using Standard Statistics to Consider Uncertainty in Industry-Based Life Cycle Inventory Databases, 10 LCA (6) 399–405 (2005)

[Tabachnik Fidell 2006]: Tabachnik BG, Fidell LS: Using Multivariate Statistics, Allyn & Bacon, Boston 2006

[Thompson 2002]: Thompson St K: Sampling, Wiley, New York 2002

[Weinrauch 1987]: Weinrauch JD: The Marketing Problem Solver, Wiley, New York 1987

[Wolter 1985]: Wolter KM: Introduction to Variance Estimation, Springer, New York 1985

Received: December 6th, 2006
Accepted: June 18th, 2007
OnlineFirst: June 19th, 2007

Annex

3.5 Calculation procedure for a representative estimate

3.5.1 Stage 1: Food markets in Berlin

For the stratified sampling in the first stage, an estimator for the average number of different types of yoghurt cups in one market is calculated as follows [Cochran 1977, pp. 91]:

$$\hat{\bar{Y}} = \sum_{h=1}^{M} \frac{N_h}{N} \bar{y}_h$$

with the variance of this estimator:

$$V(\hat{\bar{Y}}) = \frac{1}{N^2} \sum_{h=1}^{M} N_h (N_h - n_h) \frac{s_h^{\ 2}}{n_h} \text{ with } s_h^{\ 2} = \frac{1}{(n_h - 1)} \sum_{k=1}^{n_h} (y_{hk} - \bar{y}_h)^2$$

with:

N = Number of food markets in Berlin (population 1)
N_h = Number of food markets of stratum h in Berlin
M = Number of strata
n_h = Number of food markets of stratum h in the sample
$\bar{y}_h$ = average number of types of yoghurt cups in a food market in stratum h
y_{hk} = number of types of yoghurt cups in food market k in stratum h
s_h^2 = variance of the number of types of yoghurt cups in stratum h

These figures $\hat{\bar{Y}}$ are, in the first stage of the sampling, the average number of different types of yoghurt cups in one market. This average is calculated to 7.823 types of yoghurt cups per market, and the total as 9,504 yoghurt cups of different types in Berlin, for the sampling day, which is also the number of elements in the population for stage two.

3.5.1 Stage 2: Cluster sampling

For the cluster sampling in stage 2, the weight of the yoghurt cups is estimated via [Cochran 1977, pp. 250], [Kauermann 2006, pp. 89]

$$\hat{\bar{Y}}_{CL,QS} = \frac{\sum_{l=1}^{m} y_{T,l}}{\sum_{l=1}^{m} N_l} \quad \text{with the sum for the cluster:} \quad y_{T,l} = \sum_{i=1}^{N_l} y_{li}$$

The variance of this mean is estimated from:

$$\hat{V}ar(\hat{\bar{Y}}_{CL,QS}) = \frac{1}{\bar{N}^2} \frac{M-m}{M} \frac{1}{m(m-1)} \sum_{l=1}^{m} (y_{T,l} - N_l \hat{\bar{Y}}_{CL,QS})^2$$

with:

M = Number of cluster in the population
m = Number of cluster in the sample
$\bar{N}$ = Average number of yoghurt cup types in a market
N_l = Number of yoghurt cup types in market 1

3.5.3 Stage 2: Stratified sampling

For the stratified sampling in stage 2, one needs to know the share of each stratum on the population. The shares are obtained via:

$$\hat{P}_k = \sum \frac{N_h p_{hk}}{N} \quad \text{with} \quad p_{hk} = \frac{a_{hk}}{n_h}$$

with:

N = number of food markets in Berlin
N_h = number of food markets in (market) stratum h in Berlin
n_h = number of types of yoghurt cups in the sample in the market stratum h
a_{hk} = number of types of yoghurt cups in market stratum h and yoghurt cup stratum k
p_{hk} = share of yoghurt cup stratum k in market stratum h
$\hat{P}_k$ = estimator for the share of yoghurt cup stratum k in the population

12 market strata and 16 yoghurt cup strata yield 192 possible shares of yoghurt cup strata in food market strata, p_{hk}. Some of the shares are zero since not all markets offer all types of yoghurt. For all elements in one yoghurt cup stratum k, $N_k = P_k \cdot N$ holds. Consequently, for all strata, the number of elements is calculated with:

$$\begin{pmatrix} N_{k=1} \\ \vdots \\ N_{k=16} \end{pmatrix} = \begin{pmatrix} p_{1,1} & \cdots & p_{1,16} \\ \vdots & \ddots & \vdots \\ p_{12,1} & \cdots & p_{12,16} \end{pmatrix}^T \cdot \begin{pmatrix} \frac{N_{h=1}}{1215} \cdot 9504 \\ \vdots \\ \frac{N_{h=12}}{1215} \cdot 9504 \end{pmatrix}$$

A proportional sampling was applied (see above). The calculation of the estimated mean and variance for the weight of the yoghurt cups in the population is similar to calculating the estimators in the first sampling stage, for the market sampling:

$$\hat{\bar{Y}} = \sum_{k=1}^{16} \frac{N_k}{N} \bar{y}_k \quad \text{and} \quad \hat{V}ar(\hat{\bar{Y}}) = \sum_{k=1}^{16} \left(\frac{N_k}{N} \right)^2 \frac{N_k - n_k}{N_k} \frac{s_k^{\ 2}}{n_k}$$

with:

N = number of types of yoghurt cups in the population
N_k = number of types of yoghurt cups of stratum k in the population
n_k = number of types of yoghurt cups of stratum k in the sample
$\bar{y}_k$ = average weight of the types of yoghurt cups in stratum k, in the sample
s_k^2 = variance of the weight of the yoghurt cups in stratum k, in the sample

3.5.4 Stage 2: A posteriori stratification of the cluster sample

The a posteriori stratification was performed for the elements in the cluster sample by using secondary criteria for the definition of the strata in the same way as they were used in the stratified sample, in stage 1. Each element in the cluster sample was assigned to a stratum, and the share of each stratum in the population was taken from the stratified sampling. Estimators for mean and variance are calculated as in the stratified sampling procedure.

3.5.5 Calculation of uncertainty

The variances of the estimated means above are a measure for the precision of the respective estimators. They do not reflect the uncertainty of the population, which is the variation of the estimated mean. This variation, in turn, can be expressed as

$$s^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\bar{Y}})^2}{n-1}$$

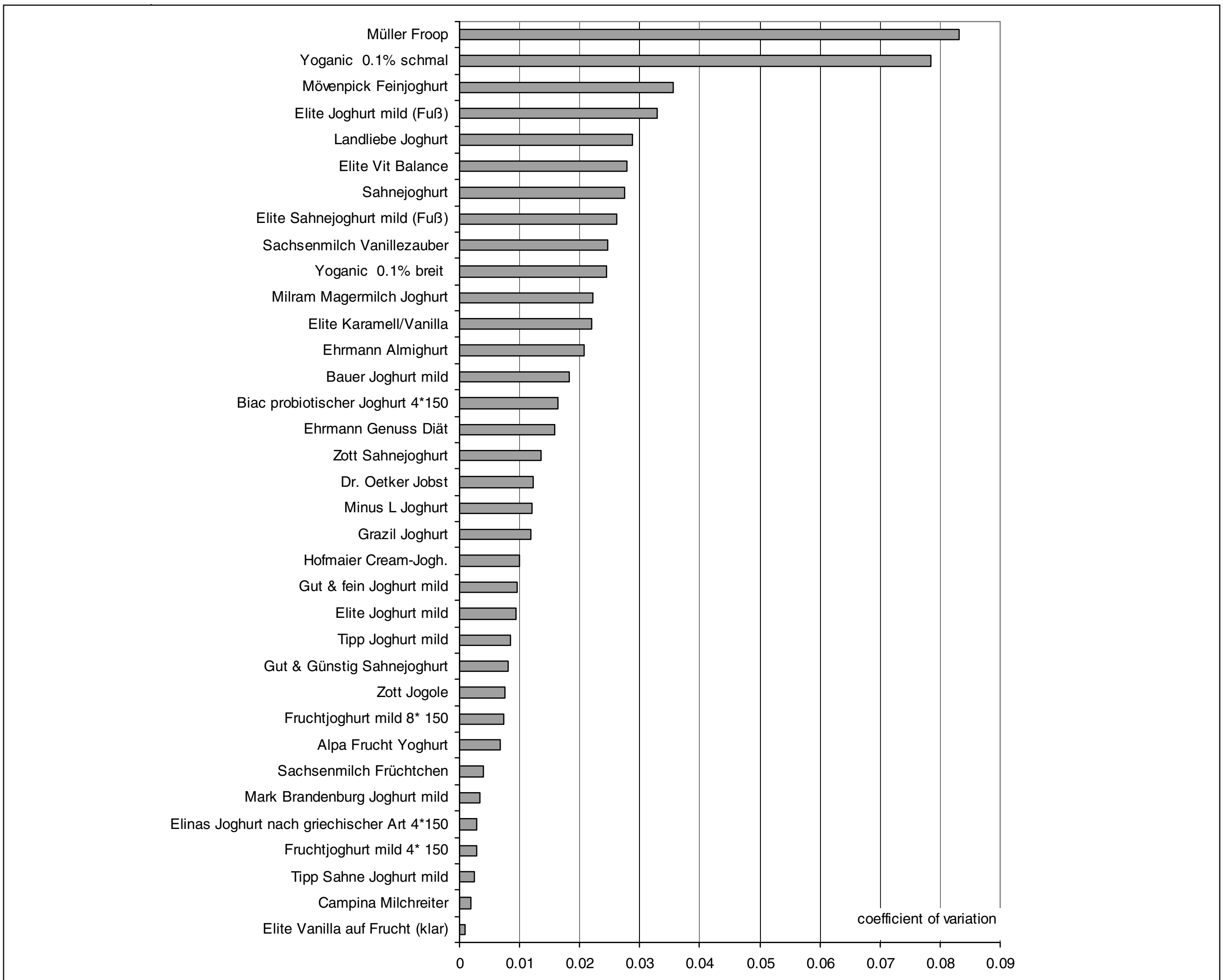

Fig. 4: (see section 3.4 'Uncertainty in samples data'). Coefficient of variation (standard deviation / mean), per type of yoghurt cup. Cups that occurred only once in the sample are excluded

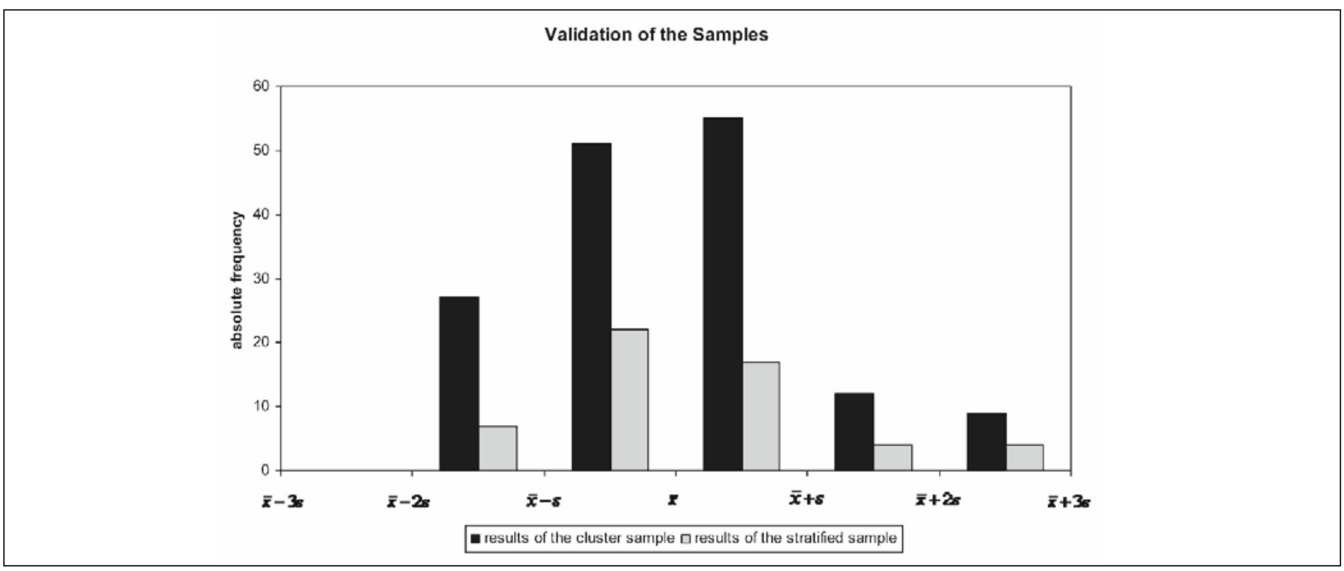

Fig. 6: (see section 3.7.2 'Sampling design'). Frequency distribution of the weight according to cluster sampling and stratified sampling s: empirical standard deviation from the cluster sample; $\overline{x}$: empirical mean from the cluster sample